



SIDE EFFECTS AND THE STRUCTURE OF DELIBERATION

BY GRANT J. ROZEBOOM

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

VOL. 9, NO. 2 | SEPTEMBER 2015

URL: WWW.JESP.ORG

COPYRIGHT © GRANT J. ROZEBOOM 2015

Side Effects and the Structure of Deliberation

Grant J. Rozeboom

THERE IS A PUZZLE ABOUT THE VERY POSSIBILITY of foreseen but unintended side effects. On the one hand, it seems clear that we can be fully rational and foresee, but not intend, bringing about some of what we take to be the important side effects of our actions. This claim, which I call “Rational Side Effects,” has a central place, not only in our moral theorizing, but also in our everyday understanding of how we relate to different aspects of our actions. On the other hand, two very plausible claims about the structure of practical deliberation jointly rule out Rational Side Effects. The first is the “Holistic Decisions” claim, which says that if an agent considers what she takes to be an important side effect of some action in the deliberation leading to her decision to perform that action, then she, if fully rational, will decide to bring about that side effect. The second is the “Decision-to-Intention” claim, which says that an agent will end up intending to do whatever she decides to do.¹ We cannot accept all three of these claims, plausible as they each seem.²

Given that we want to preserve Rational Side Effects,³ we must reject either Holistic Decisions or Decision-to-Intention. And since those latter two claims each contribute to a plausible picture of the structure of practical deliberation, resolving this puzzle will require revising that picture of practical deliberation. That is what I aim to do here. In particular, I will first explain why we should maintain Holistic Decisions, and I will propose a revised, but

¹ You might think that Decision-to-Intention receives some support from the common idea that decisions just are mental acts of forming intentions – see, e.g., Frankfurt (1988: 172); Hieronymi (2009); Holton (2009, ch. 3); and Mele (2009, ch. 1). But my account, given in section 3, replaces this common idea with an independently plausible picture of how decisions and intentions stem in different ways from the qualified intentions we hold toward our deliberative options.

² It is important to note that this puzzle is different from the much-discussed problem of “closeness”: the problem of providing an account of the morally relevant psychological difference between what an agent intends and what she merely foresees bringing about that, among other things, entails that she intends the harmful side effects that are sufficiently “close” to her intended end and means.

³ It seems to me that rejecting Rational Side Effects should be treated as a last resort in trying to solve this puzzle. One reason why, which I alluded to above, is that important moral theoretical debates, mainly about the Doctrine of Double Effect (DDE), make sense only if Rational Side Effects is true. In general, these debates take for granted that, if an agent does not intend all (or any) of the foreseen side effects of her chosen course of action, she is not thereby making some rational error. For, if that were true, opponents of the DDE could simply point out that the DDE lays out a condition on morally permissible action that can be satisfied only by an irrational frame of mind. (Note, along these lines, that the problem of closeness (n. 2) would not even be a serious problem unless we were strongly committed to Rational Side Effects.) See, for some recent discussions that illustrate this point, Bennett 1980; Fischer, Ravizza and Copp 1993; Kamm 2000; Scanlon 2008; Thomson 1999; and Quinn 1989.)

independently plausible, picture of how intentions relate to decisions, which helps to make sense of rejecting Decision-to-Intention.

My account will address the two main shortcomings of previous attempts to solve this puzzle. Gilbert Harman (1986) and Stewart Goetz (1995) reject Holistic Decisions, but, as I discuss in section 1, they fail to appreciate the steep theoretical costs of doing so: the cost of rejecting the even more compelling idea that decisions rationally serve as the conclusions of practical deliberation. Michael Bratman (1987) avoids this problem by instead rejecting Decision-to-Intention, but, as I discuss in section 2, he does not provide a revised picture of practical deliberation that adequately supports this move. Specifically, he does not provide an account of how the content of an agent's intentions reliably, but incompletely, overlaps with the content of her decisions. In section 3, I provide such an account: I argue that practical deliberation involves forming qualified intentions toward each of the options considered (an intention to carry out some action provided that one decides to do so), and I show how an agent's decisions and intentions stem in different ways from these qualified intentions of deliberation. There will thus be some reliable but incomplete overlap between her decisions and intentions: She will intend some, but not all, of what she decides to do. This account will make sense of rejecting Decision-to-Intention and help substantiate Rational Side Effects.

1.

The first piece of my account is conservative: I will defend retaining Holistic Decisions by explaining how it follows from the important and compelling idea that decisions rationally serve as the conclusions of practical deliberation.⁴ Rejecting Holistic Decisions thus requires rejecting this compelling idea. I will also show how, even though Harman and Goetz are able to deflect some of the counterintuitive implications of rejecting Holistic Decisions, they are not able to avoid this one.

Begin with the idea of a decision serving as the conclusion of some episode of practical deliberation. By "practical deliberation" I mean to designate a voluntary and typically conscious mental activity that aims at settling on what to do (Arpaly and Schroeder 2012: 211). Making a decision is how we settle on what to do at the end of deliberation. Of course, practical deliberation need not end with a decision; it may get interrupted or suspended. But if practical deliberation succeeds in achieving its characteristic aim – of settling on what to do – then it will end with a decision. When I say that decisions *rationally* function as the conclusion of practical deliberation, then, I mean that one measure along which decisions can be evaluated is how well they satisfy this aim of practical deliberation – of making an agent settled on what

⁴ This is not to say that all rational decisions result from deliberation. It is only to say that, when decisions do result from practical deliberation, they rationally function as the conclusions of that deliberation.

to do. A decision cannot entirely fail on this measure; otherwise, it would not really be a decision. But it can certainly do better or worse.

Consider, for instance, someone who realizes that he cannot both pick up his child from school this afternoon and make it to his job interview. After deliberating between these two options, he ends up deciding to go to the job interview *without* deciding not to pick up his child, not because he forgets about picking up his child, but rather because, out of discomfort with being faced with this choice, he treats not picking up his child as a matter for prediction and not decision (somewhat akin to “bad faith”).⁵ As a result, his decision fails to make him fully settled on what to do: He renews his commitment to go to the job interview, and he *believes* that he will not pick up his child, but this does not seem sufficient for settling his practical question of whether to pick up his child. His decision is thus irrational in the specific sense that I have in mind here.

How exactly does this father fail to be fully settled on what to do? Consider first what goes into being *unsettled* about what to do, which is our condition during practical deliberation. While we are still deliberating, we normally hold back from selecting any one of our deliberative options because, for each option, we are not yet willing to accept all of its costs – whether it be the cost of giving up the appealing features of the other options or the costs implied by the option itself (which, given how I am using the term “cost” here, include any of the features of the option that generate reasons against taking it). Think about the father’s ambivalence: While deliberating, he cannot yet bring himself to accept the costs of either giving up his job interview or leaving his child at school. This is what makes him unsettled about what to do.

This means that, when an agent’s decision succeeds in making her settled about what to, she thereby transitions from being not yet willing to accept the costs (broadly understood) associated with her chosen option to being willing to accept them, to what Hector-Neri Castañeda would call the “deliberate toleration” of those costs (1979: 255).⁶ I think that this suggests,

⁵ This echoes Bratman’s discussion (1987: 151-52). However, Bratman and I focus on different mistakes: Bratman focuses on the mistake of using the “Divide and Ignore” strategy in one’s deliberations, which involves noticing but then setting aside the foreseen costs of one’s options, never considering how they together bear on whether to take that option. The mistake I am discussing here grants that an agent may succeed in avoiding the Divide and Ignore strategy by viewing an action in light of all of its foreseen costs; the problem is that she may still go on to merely *predict* that she will bear those costs rather than decide to bear them. As I go on to argue, this entails that her decision fails by the rational measure of making her settled on what to do. This is important because, as I discuss below, we could accept Bratman’s claim that rational agents will avoid Divide and Ignore and still reject Holistic Decisions.

⁶ Of course, the agent may instead judge that the costs are not so steep, after all. Our decisions sometimes may help us to see the normative landscape more clearly (Holton 2010, ch. 5). But oftentimes this pattern of thought will be nothing more than a post hoc rationalization.

in turn, that her decision must in part be a decision to bear those costs. For if she only comes to believe that she will bear them, this will not be sufficient for her to become willing to bear them.

Why not? Well, think about an agent who has discovered some serious cost of her action only after deciding to carry it out. This might reasonably make her unsettled about what to do; it might reopen her question of whether to perform that action. But if believing that one's decided-upon action involved some cost were enough to resolve one's deliberative ambivalence about that cost, this should not be so. For the agent's discovery of the cost *just is* her coming to believe that her action will involve that cost. To accept the costs of her action and become fully settled on what to do, then, an agent must go beyond simply believing that she will bear those costs and include them in her decision. Now, why exactly is that so? Castañeda's phrase, "deliberate toleration," is telling.⁷ It suggests an attitude of willingness, akin to consent. Believing that one will bear certain costs does not suffice for such willingness; it is compatible with being *intolerant* of those costs – think of those who recognize a fate they are not yet willing to accept. Deciding to bear the foreseen costs of one's action requires overcoming one's intolerance of them. It involves forming an attitude of willingness toward them (even if one is not practically committed to bringing them about, and even though one still views them as costs⁸).

This sheds light on why the father's decision fails to make him fully settled on what to do. His decision leaves out the most salient cost that he associates with going to the interview, which is the cost of leaving his child at school. Even though he *believes* that going to the interview will involve leaving his child at school, his failure to include this consequence in his decision – i.e., his failure to decide both to go to the interview and to leave his child at school – thus perpetuates some of his deliberative ambivalence.

Therefore, if decisions rationally serve as the conclusions of deliberation, then they must include the unsettling aspects of chosen options. For, if a decision is fully rational, then it will make an agent settled on what to do, and if it makes him settled on what to do, then it will be a decision (in part) to bear the costs of the chosen action that he was previously not willing to bear. (Again, keep in mind that these might simply be the "opportunity costs" of giving up the attractive features of other options.) So, decisions are rationally holistic: They include the foreseen costs that contributed to an agent's unset-

⁷ I should note that, according to Castañeda, what we deliberately tolerate is included in our intention. But, in saying this, he seems more concerned with what we *intentionally* do rather than what we *intend* to do. If so, then I agree with him that we intentionally do whatever we deliberately tolerate. But, following Bratman (1987) and Di Nucci (2010), I do not think that we intentionally do only what we intend to do.

⁸ These two parenthetical qualifications are important: The first points to the space between decision and intention that I discuss below, and the second points to the difference between an agent's willingness to bear some costs and her post hoc rationalization of those costs (n. 6).

tled state in deliberation, and these will include (but will generally not be limited to⁹) the foreseen costly side effects.

At this point, you might worry that I have compromised some of the intuitive similarities between theoretical and practical deliberation. For, when we conclude theoretical deliberation and resolve our uncertainty about whether *P*, our conclusion that *P* does not seem to be subject to rational pressure to make reference to the evidence against *P* (or to the plausible *not-P* alternatives). Have we gone wrong somewhere in the above argument for Holistic Decisions, then? I do not think so, because the above argument highlights one of the respects in which theoretical and practical deliberation are importantly dissimilar. In particular, resolving theoretical uncertainty is importantly different from resolving practical uncertainty insofar as the latter but not the former generally involves overcoming one's unwillingness to bear the costs of following through on one's conclusion. This is not a general feature of reaching a theoretical conclusion, and it is the unique source of the rational pressure on practical conclusions to include foreseen costs. So it should come as no surprise that decisions (i.e., practical conclusions) are rationally holistic in this way, while the conclusions of theoretical deliberation are not.

A second worry about my argument for Holistic Decisions is that it could just as well be used to establish that intentions (and not just decisions) are rationally holistic, given that intentions, and not just decisions, are the conclusions of practical deliberation. This would undermine my eventual rejection of Decision-to-Intention. But this worry mistakenly conflates the way that decisions serve as the conclusions of deliberation and the importantly different way that intentions are conclusions of deliberation. Intentions are conclusions of practical deliberation insofar as they (like actions) are among its characteristic outputs. But they do not play the backward-looking role of concluding practical deliberation that decisions play. Rather, they play distinctively forward-looking roles in shaping future thought and action, which decisions do not play (Bratman 1987: 152-54). This suggests that the rational pressure toward holism that constrains the conclusions of practical deliberation will not apply to intentions.¹⁰

⁹ Foreseen costs will often include the expected costly means, which are not foreseen side effects. And the foreseen costs of foregoing other options may or may not be foreseen side effects; whether they are seems to depend on whether foregoing those other options is a causal upshot of one's chosen course of action, as when opening one door causes the closing of another.

¹⁰ You might notice that decisions seem to play a forward-looking role insofar as they make agents willing to bear the costs associated with their chosen options (as I discussed above). Even so, it remains that an agent's willingness (and associated decision) to bear the costs of her action will not serve as an impetus and guide to her future practical reasoning and action, as her intention does. Her willingness to bear costs will play a relatively limited background role. So I think it is still correct to understand decisions primarily in terms of their backward-looking role as conclusions of practical deliberation, and intentions primarily in terms of their forward-looking roles.

Those two worries notwithstanding, you might notice that my conclusion so far is weaker than Holistic Decisions, given that Holistic Decisions refers to *all* of what the agent takes to be the normatively significant features of her chosen action, while my conclusion so far refers only to foreseen costs. But it is a short step from here to Holistic Decisions. For in line with what Barbara Herman (1993a; 1993b) argues about an agent's "maxims," it is quite plausible to think that a rational agent will include in her decision all of the foreseen good features of her action. Herman thinks that a rational agent's maxim will include whatever features of his action he takes to generate reasons in its favor,¹¹ and the features of his action that he takes to generate reasons in its favor will often not be limited to its desirability as an end or as a means thereto; they will also include its positive foreseen side effects.¹² Setting aside the overtly Kantian notion of a maxim, Herman's line of thought thus helps fill out the intuitive idea that rational decisions are *positively* holistic. They include what an agent takes to be the good features of his action, which include its good side effects.¹³

Holistic Decisions appears to be on solid footing, then. But Harman and Goetz both reject it, suggesting that it is sufficient for rational agents to *believe* that they will bear the foreseen costs of their actions. I think they fail to appreciate, though, how Holistic Decisions follows from the idea that decisions rationally serve as the conclusions of deliberation. Examining their arguments will be instructive in highlighting this failure.

Begin with Harman's rejection of Holistic Decisions. We need to set aside Harman's own statement of this claim, which, as Bratman (1987: 149-50) points out, is too strong: Harman's version of Holistic Decisions requires not only that agents include in their decisions all of the significant foreseen side effects considered in their deliberations, but also that they include in their deliberations all of the potentially significant side effects of each of their options (Harman 1986: 99-100). I agree with Harman (and Bratman) that we should reject the second holistic requirement, but this does not bear on whether to reject the first. Focusing, then, on our more modest and plausible version of Holistic Decisions, Harman's main argument against it is that decisions rationally emulate as far as possible the "simple case" of decision-making, in which an agent decides only to bring about some end and a

¹¹ "If in willing an action an agent proceeds as she judges her action and her purpose to be good, then the maxim of action that represents her willing should contain *all* the aspects of the action and end that make them choice-worthy for her" (1993b: 222).

¹² "An agent may choose to act for an end because of her interest in it or because the end contributes to some further goal. And she may choose to act in a particular way because that action produces collateral effects also of interest to her ... If her choice is made on the basis of this rich background of value, then her maxim should include all of the aspects that determine choice-worthiness" (ibid.).

¹³ Even if you are not convinced by Herman's line of thought, you could accept a weaker version of Holistic Decisions that refers only to foreseen costs, and my main arguments here would still go through, provided that we reframed the puzzle by making Rational Side Effects also refer only to foreseen costs.

straightforward set of means for attaining that end (Harman 1986: 107-09). Even when one notices some important side effects in one's deliberations, one should relegate those side effects to one's beliefs, allowing one's decision to remain limited to a chosen end and its means (109). Why exactly should decisions be simple in this way? Harman's answer is that our decisions should be efficient in their use of our cognitive resources, and expanding their contents to include foreseen side effects will needlessly tax those resources, given that we can keep track of our actions' side effects in our beliefs whether or not we decide to bring them about (107).

My main worry about Harman's view is that he falsely supposes that the only basic, rational constraint that determines how expansive an agent's decisions should be stems from the demand for cognitive efficiency. One potential additional source of rational pressure, described by Bratman (1987: 151), is the requirement to avoid the irrational "Divide and Ignore" strategy of practical deliberation, by which an agent considers and then sets aside the costs of her options, never considering how they bear on her decision in total. But I think Harman could reply by pointing out that it is sufficient to avoid Divide and Ignore, that an agent forms a holistic belief about all of the costs of each of her options,¹⁴ which would still be more cognitively efficient than also deciding to bear those costs. Such holistic beliefs do not suffice, however, to relieve a second additional source of rational pressure: Decisions rationally serve as the conclusions of practical deliberation, and so they are constrained to have holistic contents, as I argued above. It is only if we ignore this further source of rational pressure, then, that Harman's argument against Holistic Decisions succeeds.

Goetz's case against Holistic Decisions falters for a similar reason, but he suggests a different rational constraint on the content of decisions. Rather than viewing rational decisions as cognitively cost-efficient, as Harman does, Goetz views them as responding to the reasons that can explain them. That is, he adopts the "reason-choice" principle, which says that decisions include only those features of chosen actions that agents take to provide reasons in favor of their actions and so can explain why they decide as they do (1995: 182). Thus, Goetz thinks that decisions rationally exclude negative foreseen side effects (although they may include positive foreseen side effects).¹⁵ As with Harman's argument, Goetz's line of thought supports rejecting Holistic

¹⁴ Goetz tacitly suggests as much (1995: 181).

¹⁵ This provides a response to Sarah Paul's (2011: 15) argument for thinking that practical conclusions are rationally holistic. She suggests that, so long as an agent takes some side effect into account in her deliberation about whether to perform some action, then there is no principled distinction we can draw between what is included in the agent's practical conclusion and what is relegated to her theoretical conclusion, such that the side effect is excluded from her practical conclusion. This raises a potential challenge to Harman's line of thought. But Goetz would say that we can draw such a distinction: The distinction between the considerations that are adduced in favor some action (and that can explain the relevant decision), and those that are not, and it is only if side effects fall into the former category that they will be included in the decision.

Decisions only if we make a mistakenly narrow assumption about the rational pressures that govern how expansive our decisions should be – in Goetz’s case, only if we assume that these pressures all stem from the fact that decisions answer to the reasons that can explain them. This assumption is false because, again, decisions rationally serve as the conclusions of practical deliberation and so answer more broadly to the considerations taken into account in deliberation.

Thus, I think that Holistic Decisions stands up to scrutiny. It remains firmly supported by the fact that decisions rationally serve as the conclusions of practical deliberation. This means that we should look elsewhere to solve our puzzle.

2.

Given that we want to maintain Rational Side Effects, our only other option is to reject Decision-to-Intention. This is the route that I think we should take, but it starts with an uphill climb. For in order to plausibly reject Decision-to-Intention and maintain Rational Side Effects, we need to provide a revised account of how intentions relate to decisions, such that a rational agent need not intend to do whatever she decides to do, and yet there is still *some* reliable overlap between the contents of her intentions and decisions. Such an account is needed because, as Harman points out, once we reject Decision-to-Intention “it is extremely difficult to see what the connection could be between one’s intentions and one’s practical conclusions” (1986: 98). In this section, I explore what is required to meet this challenge by considering how Bratman’s attempt falls short.

Bratman begins with the idea that practical deliberation involves the construction of a holistic “deliberative scenario” associated with each option (143-44). As we proceed in deliberation, noting both attractive and aversive features of our options, we expand our conceptions of these options, which often begin as simple (and perhaps incomplete) chains of means to some end but grow to include further means and side effects. Bratman claims that our decision will be a choice in favor of one of these scenarios as a whole, and not just in favor of some part of the scenario, because, as I mentioned above, this is how he thinks that we can avoid the irrational Divide and Ignore strategy of practical deliberation. And he argues that our resulting intention will not always be so expansive, because we will sometimes have relevant “self-governing” intentions that are “functionally incompatible” with intending certain of the side effects of our chosen action and that will thus filter those side effects from our intention (159-60). That is, we will sometimes maintain personal policies that would be in tension with intending to bring about certain of the side effects contained in our chosen deliberative scenario, and, as a result, we will rationally not intend to bring about those side effects (unless

we give up our personal policies).¹⁶ So, this is how Bratman thinks that we can make sense of how intentions relate to decisions even though Decision-to-Intention is false: Both decisions and intentions concern the holistic scenarios that we construct in practical deliberation, but only our decisions will generally concern the whole scenario, because our intentions are sometimes constrained by self-governing intentions that hold us back from intending the side effects in those scenarios.

While I of course agree with Bratman that decisions are rationally holistic, and while it seems correct to think that self-governing intentions do sometimes filter side effects from our intentions, I do not think that Bratman's account meets the challenge facing views that reject Decision-to-Intention. The first problem was mentioned earlier: His rationale for Holistic Decisions – which he understands in terms of choosing the whole deliberative scenario – does not actually entail that we are generally rationally required to make holistic decisions. For his rationale rests on the idea that we should avoid the Divide and Ignore strategy of practical deliberation, and (as Bratman himself acknowledges) making a holistic decision is only one way of avoiding Divide and Ignore. Another good way to avoid it is to simply believe that one's course of action will have all of the features included in the deliberative scenario. This problem resurfaces here, because it means that, in the wake of rejecting Decision-to-Intention, we do not yet have an explanation of why there will always be some overlap between an agent's decision and intention. For all Bratman has said, a rational agent's decision and intention could concern entirely different parts of the same deliberative scenario, provided that she believes that she will bring about the whole scenario. For instance, perhaps her intention simply concerns the end *E* and the chain of means leading to it, and her decision simply concerns one of the desirable side effects of pursuing *E*. She will avoid Divide and Ignore by forming a holistic belief about what she is doing in pursuing *E*, and she will satisfy the relevant rational constraints that come to bear on her intention.¹⁷ Yet, there is no overlap between her decision and intention.

Suppose we grant that Bratman could solve this problem, perhaps by incorporating the argument for Holistic Decisions given in section 2. (His account is certainly compatible with my view of how decisions rationally serve as the conclusions of practical deliberation.) Even so, a second problem arises. If an agent making a decision does not have the right kind of self-governing intention – one that would be incompatible with intending certain of the foreseen side effects of her chosen course of action – then, on Bratman's account, she should just as well intend to realize the whole deliberative

¹⁶ Bratman develops the idea of a self-governing policy in detail in later work – see esp. Bratman 2007. For our purposes here, the details of Bratman's later account will not matter.

¹⁷ This glides over some complexities about the rational norms that govern intention that are important but not relevant here. See, e.g., Bratman 2009; Broome 2013; Kolodny 2007; and Raz 2005.

scenario.¹⁸ That is, given that there is no rational pressure against intending the whole scenario (since, by hypothesis, she does not maintain any counter-vailing self-governing intentions), it seems sensible for her to intend to realize the whole package. Of course, she is not required to do so as a matter of means-end rationality. But, given that she is faced with the whole scenario, this does not explain why she would intend to realize only part of it rather than the whole thing. You might think that, following Harman, considerations of cognitive efficiency would tell in favor of intending the smallest part of the scenario that she is required by means-end rationality to intend. But I think the opposite may be true. The scenario is presented as a whole, and so, if the agent is only to intend part of it, then cognitive resources will need to be expended in selecting the part that she will intend. This would suggest that, on Bratman's account, the agent will form an intention to realize the whole deliberative scenario (again, given that she does not have a self-governing intention that could filter out side effects from her intention). This is a problem because, again, in order to plausibly reject Decision-to-Intention and maintain Rational Side Effects, one of the things that we need from our revised account of practical deliberation is an explanation of why a rational agent's intentions can exclude the foreseen side effects that are included in her decisions.

Now, you might think there is no real problem here, because all we need to reject Decision-to-Intention and maintain Rational Side Effects is an explanation of why a rational agent will *sometimes* exclude from her intentions the side effects that are included in her decisions, and Bratman's account gives us that much. But while this may suffice for rejecting Decision-to-Intention, it does not give us what we need to substantiate Rational Side Effects. For it seems clear that Rational Side Effects (or, at least, the intuitive thought behind it) does not concern only those cases in which agents have the right self-governing intentions to filter foreseen side effects from their intentions. It is meant to capture a more general thought about rational agency: that, in normal circumstances – and, we should now add, regardless of whether she maintains pertinent self-governing intentions – a fully rational agent may exclude foreseen side effects from her intentions.

The two problems facing Bratman's account bring into sharper focus the basic tasks of an adequate revised account of how intentions relate to decisions: (1) it must explain why there is always *some* overlap between the intention and decision resulting from an agent's practical deliberation, and (2) it must explain why this overlap can normally, and not just in special cases, be incomplete. Without (1), our account will be left without an answer to Harman's pressing question: What *is* the connection between our decisions

¹⁸ Or even if she does now maintain some relevant self-governing intention, you might think that she could restore rational order in her mind simply by giving up that self-governing intention and still intending to bring about the side effects. But Bratman's (2007) later views about self-governance tell against this strategy, given the role that self-governing intentions play in constituting an agent's identity across time.

and intentions, if an agent does not intend to do whatever she decides to do? And without (2), we will be unable to make sense of Rational Side Effects (even if we succeed in rejecting Decision-to-Intention), because Rational Side Effects does not concern special cases, but rather captures a more general truth about the intentions of rational agents.

3.

What we are still looking for, then, is an account that completes the tasks of (1) and (2). That is what I attempt to develop in this section. My account rests on the idea that practical deliberation involves forming a qualified (or conditional) intention to carry out each of one's deliberative options: an intention to carry out some action A , provided that one decides to A . Below, I first explain why we should think that practical deliberation involves qualified intentions of this sort, and then I show how this account of practical deliberation helps to generate the explanations required by (1) and (2) and, thus, meets the challenge of rejecting Decision-to-Intention and securing Rational Side Effects.

Consider an agent in the midst of practical deliberation. She is considering whether to ask her supervisor if she can take off the Fourth of July in order to watch her son march in the local Independence Day parade. Notice that, in considering this deliberative option, she has assigned her prospective action – of asking her supervisor if she can have the holiday off work – the end of watching her son march in the parade. This manifests a general requirement of deliberation: In order to deliberate about whether to carry out some action A , one must assign some end E to A . For if this were not so, then deliberating agents could not undertake the various activities that are constitutive of practical deliberation, such as weighing reasons and applying general principles. Many of our reasons derive from the ends our actions serve, and we often cannot apply general principles to our prospective actions without first considering their end(s). So, practical deliberation requires that we assign ends to our prospective actions.¹⁹

What does it take to fulfill this requirement? It is not enough (although it may be necessary) that an agent believes that, should she decide to A , her A -ing will promote E . For she might believe this about many ends, none of which are the end(s) she assigns to A . For instance, our agent may notice that her action of asking to take off the holiday will serve the end of distracting her supervisor from the poor performance of her colleagues. In noticing this, she need not assign this end to her prospective action. It is also not enough that, in addition to believing that her prospective A -ing will promote E , she desires that this be so. For she might have this desire-belief pairing toward many possible ends, none of which are the end(s) she assigns to her option

¹⁹ Note that sometimes the end is simply the action itself. But even when this is so, we usually build in some idea of the more proximal means that will be needed in order to execute the action.

of *A*-ing. Imagine, for instance, that our agent also desires that her prospective action of asking to take off the holiday will distract her supervisor from the poor performance of her colleagues. Still, this need not be the end(s) she assigns to this option. She might view this simply as a happy side effect of her option. So, not even both believing and desiring that one's prospective *A*-ing will promote *E* is sufficient for assigning *E* to *A* in one's deliberation about whether to *A*.

This suggests that one must intend for one's prospective *A*-ing to promote *E*. The reason why believing and/or desiring does not suffice, it seems, is that these states of mind do not (neither together nor separately) realize how deliberating agents will be committed, albeit provisionally, to having their prospective actions serve certain ends. Indeed, the formation of such provisional commitments seems to signal the initiation of deliberation; this is an important part of how an agent transitions from idly wondering about *A*-ing, or not even thinking about *A*-ing at all, to deliberating about whether to *A*. And the natural way to understand these provisional commitments is as provisional intentions.²⁰

But how? What exactly goes into maintaining these provisional intentions? They are provisional insofar as they are qualified by the outcome of deliberation, i.e., on whether the agent decides to take the option. And they are qualified by this outcome, not in the external sense that whether an agent has the intention depends on the outcome (for we have already accepted that the intention is present during her deliberation), but rather in the internal sense that what the agent intends is qualified. That is, what the agent intends to do, which is to carry out *A* in order to promote *E*, is explicitly conditioned on some outcome, which here is the outcome of her deciding to *A*. So, we can represent her intention as an intention to [*A* in order to promote *E*, provided that she decides (or "I decide"²¹) to *A*].

Following Luca Ferrero's (2009) account of qualified (what he calls "conditional") intentions, it is important to distinguish how this intention is internally qualified from two other ways that intentions may be thought to be internally qualified.²² First, almost all of an agent's intentions are internally qualified in the sense that, were she to believe that certain outcomes have

²⁰ One distinctive mark of practical deliberation is that it involves a commitment, or intention, to settle the matter of what to do – see again Arpaly and Schroeder (2012: 211) and also Smith (2009: 3-4). I am now suggesting that it also involves a provisional commitment to each of one's options, which cannot be straightforwardly derived from the overarching commitment to settle the matter of what to do.

²¹ It might be more accurate to represent the content first-personally, as from the agent's point of view, but for the sake of keeping a consistent voice in writing, I will continue using third-person pronouns.

²² My main aim here is to use the notion of a qualified intention to understand some otherwise puzzling facts about deliberation. I do not aim to introduce or provide a full explication of this notion; for that, I rely on Ferrero's (2009) account. But to spell out this notion a bit more, I discuss below how the rational norm of means-end coherence applies in a distinctively limited way to qualified intentions (n. 25 and n. 30).

transpired, she would give up the course of action that she now intends to carry out. She does not intend to carry out her actions in all conceivable futures. But usually these conditions do not figure *explicitly* into the content of her intentions. They are not held before her mind as she is guided by the intention, as they are when deliberating agents are guided by the qualified intentions they hold toward their options. (This might have to do with the fact that deliberating agents take themselves to have control over the conditioning outcome, i.e., over what they decide to do.) Second, a deliberating agent's intentions are not internally qualified in virtue of having a material conditional serve as their content. For a deliberating agent is not guided by her qualified intention toward the option of *A*-ing if she makes (what would be) the antecedent of her intention false by failing to decide to *A*.²³ The condition of her intention thus does not function as the antecedent of an intended material conditional, but rather plays the role, as Ferrero (2009: 205) puts it, "of ... 'setting the stage' for the performance" of what she intends to do. She is "aiming at [carrying out her action] against the backdrop of the obtaining of [the condition]." A deliberating agent, in particular, aims at *A*-ing so as to achieve *E* against the as of yet uncertain backdrop of her deciding to *A*.²⁴

Let us say, then, that an agent who is deliberating about whether to *A* intends to [*A* in order to promote *E*, provided that she decides to *A*], keeping these two clarifications in mind.²⁵ How is this supposed to help us meet the challenge of rejecting Decision-to-Intention (while supporting Rational Side Effects) by providing, as (1) and (2) require, an explanation of why there can normally be some incomplete overlap between an agent's decisions and

²³ See Ferrero (2009: 703-05).

²⁴ You might worry that this claim implies that practical deliberation is generally higher order in ways that it clearly is not – see Broome (2013, ch. 12). In particular, the qualified intentions of deliberation explicitly refer to an agent's prospective decisions, and not just to her prospective actions and/or the reasons for and against them. But this would make her practical deliberation implausibly higher order only if these qualified intentions supplied premises in her practical reasoning, which they do not. Thanks to Michael Bratman for bringing this concern to my attention.

²⁵ Further support for this claim may be found by reflecting on the limited (see n. 30) demands of means-end rationality that come to bear on a deliberating agent. Suppose, for instance, that our agent who is deliberating about whether to ask to take off the Fourth of July must make her request to her supervisor in person by 5:00 pm today, it is now 3:00 pm, and it will take her two hours (using public transit, as she must) to travel to the office. It seems that, on pain of irrationality, she must intend to now begin traveling to the office or else cease deliberating about whether to make the request of her supervisor. This would be so only if she now intends (albeit qualifiedly) to make the request of her supervisor. You might object by suggesting that we can explain this irrationality simply in terms of how an agent must believe that her deliberative options are possible – see Nelkin (2004) for some relevant discussion. But this explanation is insufficient, because the agent could falsely believe for good reasons that she intends to now go to the office and thereby alleviate whatever irrationality stems from the rational requirement to believe in the possibility of her options. Yet, given that she does not actually intend to now go to the office, she will remain irrational, and this remainder of irrationality could only be due to her qualified intention to make the request of her supervisor. Here, I am helped by the argument made by Bratman (2009: 428).

intentions? The key is to see how a rational agent's decisions and intentions stem in different ways from the qualified intentions of her deliberation.

Her decisions will contain the conceptions of her chosen options that are given by these qualified intentions, because these are the basic conceptions of her options that she considers in her deliberation. Following what I said above, each of these conceptions will consist in her basic idea of some action serving the end that she has assigned to it. If her decision is going to serve as the conclusion of her deliberation and holistically include the foreseen side effects of her chosen option, then it must also include this basic conception of her option to which those side effects are attached. We can represent this by saying that if her deliberation involved an intention to [*A* in order to promote *E*, provided that she decides to *A*], then her decision will in part be a decision to [*A* in order to promote *E*]. Now, you may notice that, so far, I have been representing the condition of an agent's qualified intention simply as a decision to *A*. But in light of what I just said, along with the main argument of section 1, it is clear that this decision will be much more than a decision to *A* – it will be a decision to [*A* in order to promote *E*], along with the salient good- and bad-making features of *A*, including the foreseen side effects of *A*-ing. This complication does not present any problems for my view, however: We can plausibly expand our description of the decision in the content of the qualified intention, and/or we can accept that a decision to [*A* in order to promote *E*, along with the important features of *A*-ing] entails deciding to *A*.

An agent's intentions will also stem from the qualified intentions of her deliberation, but in a more direct way: They will be the unqualified versions of her qualified intentions. How so? Whenever an agent maintains an internally qualified intention to [*A*, provided that *C* obtains] and realizes that the qualifying condition *C* on her intention obtains, she will come to unqualifiedly intend to *A*.²⁶ For, as was suggested above, intending to [*A*, should *C* obtain] amounts to aiming at *A*-ing against the as-of-yet-uncertain backdrop of *C* obtaining, i.e., aiming at *A*-ing in the limited range of foreseeable circumstances in which *C* obtains. So, when she believes that *C* obtains, she will be aiming at *A*-ing in the circumstances in which she finds herself. She will be unqualifiedly intending to *A*. This is what happens when an agent, who during deliberation qualifiedly intends to [*A* in order to promote *E*, provided that she decides to *A*], decides in favor of this option. For decisions are characteristically self-conscious mental acts, such that when an agent decides to do something, she realizes that she has made that decision. So, the agent will be aiming to [*A* in order to promote *E*] in the very circumstances in which she finds herself, in which she has decided to [*A* in order to promote *E*]. She will thus be unqualifiedly intending to [*A* in order to promote *E*].

²⁶ Just in the sense that most of our everyday intentions can be unqualified: There are qualifications, but they are implicit. They do not figure explicitly into the content of the intention.

This account of how decisions and intentions stem in different ways from the qualified intentions of deliberation helps to explain why (1) there is always some overlap between a rational agent's decisions and intentions, and why (2) this overlap is (and thus can be²⁷) normally incomplete. Begin with (1). On my picture, if a rational agent considers the option [A in order to promote E] and she decides in favor of that option, then she will both decide to [A in order to promote E] and intend to [A in order to promote E]. For, given that her deliberation involves the qualified intention to [A in order to promote E , provided that I decide to A], then, once she decides in favor of that option, she will come to unqualifiedly intend to [A in order to promote E]. And her decision will include [A in order to promote E], since this is the basic conception of her option that she uses in her deliberation. So we have an explanation of why there will reliably be some overlap between what a rational agent decides to do and what she intends to do.

But, turning to (2), this overlap will normally be incomplete, because her intention will be limited to what was included in her corresponding qualified intention of deliberation, and her decision will not be limited in this way. For it will also include the important foreseen side effects of her chosen action. This raises the important question, though, of why the qualified intentions of her deliberation will not expand to include foreseen side effects. They will not, first, because they are not subject to the rational pressure toward holism, given that they do not serve as the conclusions of practical deliberation. And they will not expand to include foreseen side effects in the absence of such pressure, because, second, doing so would place unnecessary burdens on the cognitive demands of her practical deliberations.²⁸ This is a point where the rational demand for cognitive efficiency does tell against holism. Given that a rational agent's qualified intentions of deliberation do not initially include foreseen side effects, and given that she can keep track of these side effects by other means (namely, with her beliefs and decisions), then she will not include these side effects in her qualified intentions.²⁹ This explains why, then, a rational agent will normally intend to do less than what she decides to do – she will not intend to bring about the side effects that are included in

²⁷ My account supports a somewhat stronger conclusion than the one we need: It explains why rational agents generally *will not* intend foreseen side effects, which entails but goes beyond the claim that they generally *may* exclude these side effects from their intentions. But I do not think it is implausibly stronger.

²⁸ This is not to deny that, along the lines that Bratman suggests, our personal policies and commitments can also play a role in filtering foreseen side effects from our intentions and, more specifically, from the qualified intentions of our deliberations. It is just to explain why, even when no such commitments come to bear, we generally will not include side effects in our qualified intentions.

²⁹ This is not to make Harman's mistake (discussed in section 1) of ignoring sources of rational pressure beyond cognitive efficiency. As I mentioned above (n. 25), the qualified intentions of deliberation will also be subject to a rational norm of means-end coherence, which calls for these qualified intentions to include some of the means that the agent discerns during deliberation.

her decision – even though there will be some significant overlap between her decisions and intentions.

Why exactly, though, would it be cognitively costly for an agent to include foreseen side effects in the qualified intentions of her deliberation? If she were to include them, then she would need, among other things, to keep track of the necessary means for bringing about these side effects, given the rational pressures of means-end coherence that come to bear on qualified intentions.³⁰ (As it is, she does not think that there are necessary means beyond those already contained in her deliberative options. But she would need to be responsive to evidence suggesting that there were further means she needed to take.) This would not be so if, for each of her options, she simply believed that she would bring about its side effects if she took it.³¹

My account of how decisions and intentions stem in different ways from the qualified intentions of deliberation thus gives us what we need to reject Decision-to-Intention and maintain Rational Side Effects. It helps us to understand why there is reliably some incomplete overlap between a rational agent's intentions and decisions, in three steps. First, looking back to the argument for Holistic Decisions, decisions rationally include foreseen side effects because of what it takes for decisions to serve as the conclusions of practical deliberation. Second, intentions are limited to what is included in the qualified intentions of deliberation, and these qualified intentions tend to exclude foreseen side effects (for reasons of cognitive efficiency). Third, decisions include, but (given the first step) are not limited to, what goes into the qualified intentions of deliberation.

You might wonder at this point how exactly my view improves upon Bratman's. Granted, it is different; but what exactly are the resources it has that Bratman's view lacks? The most important difference is the idea that the relatively narrow qualified intentions of deliberation, and not the holistic deliberative scenarios that Bratman describes, provide the basis for a rational agent's intentions. For this is what gives us grounds for concluding that her intentions can normally exclude the foreseen side effects that are included in her decisions. Rather than leave us wondering why her intentions will con-

³⁰ These rational pressures are not as expansive as those that apply to unqualified intentions, though. Specifically, while both qualified and unqualified intentions are subject to some norm of means-end coherence (n. 25), this norm seems limited to *imminently* necessary means in the case of qualified intentions (what the agent sees as means that she must take now), while it is not limited in this way for unqualified intentions – it covers all, or more, of what the agent takes to be necessary means (Ferrero 2009: 707, 711). Even so, note that it still makes sense to think of qualified intentions as genuine intentions. In addition to the (limited version of the) norm of means-end coherence, they are subject to the rational norms of consistency and stability that are characteristic of intention – see *ibid.* (711-12). Thanks to an anonymous referee for pressing me to clarify these points.

³¹ Nor need she do so if, after deciding in favor this option, she is merely willing to bring about the side effects (and does not intend to do so). The attitude of willingness does not put us under rational pressure to intend the means to whatever we are willing to bring about, as intention does.

cern only some part of her holistic deliberative scenarios, my view explains why her intentions will be limited to the basic conceptions of her options contained in her qualified intentions.³²

4.

To conclude: A puzzle arises from the common claim (Rational Side Effects) that we can rationally exclude foreseen side effects from our intentions, and this puzzle forces us to rethink two of our basic ideas about the structure of practical deliberation: Holistic Decisions, which says that decisions rationally must include all important foreseen side effects, and Decision-to-Intention, which says that an agent intends to do whatever she decides to do. For we cannot maintain both of these ideas and accept that agents can rationally exclude foreseen side effects from their intentions. I have argued that we should keep Holistic Decisions and reject Decision-to-Intention.

In particular, I first showed how Holistic Decisions is supported by the compelling idea that decisions rationally serve as the conclusions of practical deliberation. To become fully settled on what to do, an agent will decide (and so become willing) to bring about foreseen side effects. To make sense of rejecting Decision-to-Intention, I then provided an account of how our intentions derive directly from the qualified intentions that we maintain during practical deliberation. These qualified intentions will normally exclude foreseen side effects, and so our intentions will exclude them, as well. My account thus solves the puzzle by providing a plausible explanation of why a rational agent will normally not intend to bring about the side effects that she decides to bring about.

Note that, while my account was initially motivated by the desire to solve the puzzle about foreseen but unintended side effects, it rests on independent grounds – some observations, first, about how decisions rationally serve as the conclusions of practical deliberation and, second, about how deliberating agents assign ends to their prospective actions. In this way, confronting the puzzle has furthered our theorizing about the structure of practical deliberation.³³

Grant J. Rozeboom
Stanford University
Department of Philosophy
rozeboom@stanford.edu

³² A further difference between our views, which I discussed earlier, is that my argument for Holistic Decisions explains why holistic beliefs are insufficient to satisfy the rational demand of holism, and Bratman's account does not.

³³ Many thanks to Samuel Asarnow, Michael Bratman, John Fischer, Pamela Hieronymi, Sarah Paul and Tamar Schapiro for generous comments on earlier drafts of this paper, and to two anonymous referees for the *Journal of Ethics & Social Philosophy* for their helpful suggestions.

References

- Arpaly, N. and Schroeder, T. (2012) "Deliberation and Acting for Reasons," *The Philosophical Review* 121(2): 209-39.
- Bennett, J. (1981) "Morality and Consequences," in *The Tanner Lectures on Human Values* (Vol. 2, ed. S. McMurrin), Salt Lake City: University of Utah Press, pp. 45-116.
- Bratman, M. (2009) "Intention, Practical Rationality, and Self-Governance," *Ethics* 119(2): 411-43.
- _____. (2007) *Structures of Agency*, New York: Oxford University Press.
- _____. (1987) *Intention, Plans, and Practical Reason*, Cambridge, Mass.: Harvard University Press.
- Broome, J. (2013) *Rationality Through Reasoning*, Malden, Mass.: Wiley-Blackwell.
- Castañeda, H. (1979) "Intensionality and Identity in Human Action and Philosophical Method," *Noûs* 13(2): 235-60.
- Di Nucci, E. (2010) "Rational Constraints and the Simple View," *Analysis* 70(3): 481-86.
- Ferrero, L. (2009) "Conditional Intentions," *Noûs* 43(4): 700-41.
- Fischer, J., Ravizza, M. and Copp, D. (1993) "Quinn on Double Effect: The Problem of 'Closeness,'" *Ethics* 103(4): 707-25.
- Frankfurt, H. (1988) "Identification and Wholeheartedness," in *The Importance of What We Care About*, New York: Cambridge University Press, pp. 159-76.
- Goetz, S. (1995) "The Choice-Intention Principle," *American Philosophical Quarterly* 32(2): 177-85.
- Harman, G. (1986) *Change in View*, Cambridge, Mass.: MIT Press.
- Herman, B. (1993a) "The Practice of Moral Judgment," in *The Practice of Moral Judgment*, Cambridge, Mass.: Harvard University Press, pp. 73-93.
- _____. (1993b) "Leaving Deontology Behind," in *The Practice of Moral Judgment*, Cambridge, Mass.: Harvard University Press, pp. 208-42.
- Hieronymi, P. (2009) "The Will as Reason," *Philosophical Perspectives* 23: 201-20.
- Holton, R. (2009) *Willing, Wanting, and Waiting*, Oxford: Oxford University Press.
- Kamm, F. M. (2000) "The Doctrine of Triple Effect and Why a Rational Agent Need Not Intend the Means to His End," *Aristotelian Society Supplemental Volume* 74(1): 21-39.
- Kolodny, N. (2007) "How Does Coherence Matter?" *Proceedings of the Aristotelian Society* 108(3): 229-63.
- Mele, A. (2009) *Effective Intentions: The Power of Conscious Will*, Oxford: Oxford University Press.
- Nelkin, D. (2004) "Deliberative Alternatives," *Philosophical Topics* 32(1&2): 215-40.
- Paul, S. (2011) "Deviant Formal Causation," *Journal of Ethics and Social Philosophy* 5(3): 1-23.
- Quinn, W. S. (1989) "Actions, Intentions, and Consequences: The Doctrine of Double Effect," *Philosophy and Public Affairs* 18(4): 334-51.
- Raz, J. (2005) "The Myth of Instrumental Rationality," *Journal of Ethics and Social Philosophy* 1(1): 1-28.
- Scanlon, T. M. (2008) *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge, Mass.: Harvard.
- Smith, M. (2009) "Practical Imagination and its Limits," *Philosophers' Imprint* 10(3): 1-20.
- Thomson, J. J. (1999) "Physician-Assisted Suicide: Two Moral Arguments," *Ethics* 109(3): 497-518.